ABSTRACT

Methodological problems applying logistic regression for risk calculations from accident data

It is a common practice to apply logistic regression when estimating risk from accident data for hit pedestrians, i.e. to create pedestrian fatality risk curves. This is though problematic since pedestrian accidents with required data available are limited in number and of great complexity in their nature. This results in that the data material usually includes relative high number of accidents for the lower spectrum of the speed range, i.e. the urban speed environment 30-50 km/h, but only a handful of accidents for the higher end of the speed spectrum.

The aim of this presentation is to analyze how this problem affects the results of logistic regression from this type of databases. It will though also look at more fundamental aspects such as if the logistic curve (an S curve) is logical for describing the fatality risk for hit pedestrians and other mathematical properties relevant for applying and interpreting the quality of the resulting fatality risk curve.

Höskuldur Rúnar Guðjónsson
Department of Technology and Society
Lund University
BOX 118
221 00 LUND, SWEDEN
+46 (0) 46 222 91 98
hoskuldur.gudjonsson@tft.lth.se